社団法人 情報処理学会　研究報告
IPSJ SIG Technical Report

2004－DBS－133　(8)
2004－ＦＩ－ 75　(8)
2004／5／14

社会的関係を用いた評価コメントの自動要約方法とその有効性の検証

大野華子[†]　楠村幸貴[†]
土方嘉徳[†]　西田正吾[†]

ネットオークションには落札者が記述した出品者に対する評価コメントがあり，ユーザはこれを参考にしてどの出品者から商品を購入するかを選択する．しかし，これらの評価コメントは大量に存在する上に儀礼的な文を多く含むため，ユーザが複数の出品者を比較するのは大変な作業である．本研究ではこの問題に対し，出品者評価コメントをネットオークション上の社会的関係を用いて要約する手法（Social Summarization 法）を提案する．本稿では，出品者評価コメントについて行った調査とその結果から考案した Social Summarization 法の概要について述べる．また，Social Summarization 法を実装した要約システムと Social Summarization 法の有効性の評価について述べる．

Evaluation of Summarization Method for Feedback Comments using Social Relationship

Hanako Ohno,[†] Yukitaka Kusumura,[†]
Yoshinori Hijikata[†] and Shogo Nishida[†]

Bidders on net auctions write feedback comments to the sellers from whom the bidders have bought the items. Other bidders read them to determine which item to bid for. In this research, we aim at supporting bidders by summarizing the comments. First, we examine feedback comments on net auctions and show the result of the examination. After that, we propose a method called Social Summarization method, which uses social relationships on net auctions for summarizing feedback comments. And we implement a system based on our method and finally evaluate its effectivity.

1. はじめに

近年，電子商取引の中でもネットオークションが賑わいを見せている．通常，ユーザがネットオークションで商品を購入する場合は，キーワードなどで商品を検索し，入札する商品を決定するために商品の紹介文と出品者評価コメントを読む．ここで商品の紹介文とは，出品者が商品の特徴や性能について記述した文章である．また出品者評価コメントとは，落札者が記述した出品者に対する評価についての文章である．出品者評価コメントは，入札者が入札する商品を決定する際の参考になるように，出品者ごとにまとめられている．しかし，出品者一人あたりの評価コメントの数は非常に多く，出品者によっては数千個に及ぶ．また，評価コメント中にはお礼や挨拶などの儀礼的な文がたくさん含まれており，その中に埋もれている出品者に関する有益な情報を抽出するのは大変困難である．

本研究では，落札者ごとにその人の記述した出品者評価コメントを調べ，注目している出品者に対する評価コメントとその他の出品者に対する評価コメントを比較して要約する手法（Social Summarization 法）を提案する．本手法では，注目している出品者にのみ存在する記述と注目している出品者にのみ存在しない記述を抽出して要約とする．これにより，複数の出品者を比較するために有益な情報を，まとめてユーザに提示することができる．ここで記述とは，文章中の意味的な単位で文や文中の節のことを指す．

本研究のアプローチは，どのように要約すれば精度の高い要約を作成できるかを知るために，まず出品者評価コメントの調査を行う．次に，その調査結果に基づき，要約手法を考案する．そして，この要約手法を使った要約システムを実装し，要約手法の有効性を評価する．本論文の以降の構成は次のようになっている．2 章で関連研究について述べ，本研究との違いを明確にする．3 章ではネットオークションの出品者評価コメントの調査について述べる．4 章では調査結果をもとに考案した Social Summarization 法の概要について述べる．5 章では実装した要約システムについて述べる．6 章で，Social Summarization 法の有効性の評価について述べる．そして，7 章でまとめと今後の課題について述べる．

† 大阪大学大学院基礎工学研究科
Graduate School of Engineering Science, Osaka University

表 1 インフォーマルな文章を要約する手法
Table 1 Comparison with the existing systems which summarize informal texts.

システム名	対象	抽出するもの	抽出または削除する情報
MHC 1)	メール，Web上の予定情報	予定の件名，日付	固定
fj.meetings 要約システム 2)	会議案内のネットニュース	会議名，日時，場所	固定
fj.wanted 要約システム 3)	求人・売買などのネットニュース	求人・売買の種類とその内容をよく表す文	固定
CIKLE 4)	Web掲示板上のメッセージ	話題に関する内容を表す文	固定
本研究	ネットオークションの出品者評価コメント	出品者の特徴をよく表す文	適応的に選択

2. 関連研究

本研究の関連研究として，インフォーマルな文章を要約する手法について取り上げる．インフォーマルな文章とは，新聞やビジネス文書などで見られる，整形された，かつ正しい書き言葉で書かれた文章とは異なり，掲示板やチャットなどで見られる，不特定多数のユーザにより形式張らずにやり取りされるメッセージのことである．要約する手法の具体的な例として，乃村らの MHC[1)]，佐藤らのネットニュースのダイジェスト自動生成システム[2),3)]，梅木らの CIKLE[4)] を挙げる（表 1 参照）．

MHC で用いられる手法では，予定情報を含むメールや Web ページから，予定の件名と日付を抽出する．予定の件名はメールや Web ページのタイトルから抽出する．また，日付は，「月」，「日」，「/」などの文字や記号を基にして取り出している．

佐藤らのシステムは，学会の案内を流すニュースグループ（fj.meetings）と，求人や売買などの記事が投稿されるニュースグループ（fj.wanted）を対象としたものがある．それぞれ，抽出する対象に特有な言語表現のパターン（開催日ならば “∗ 年 ∗ 月 ∗ 日” など）と，行の開始位置や箇条書などのテキストのレイアウトを手がかりとして抽出する．

CIKLE で用いられる手法では，メールの投稿によって作成される Web 掲示板上のメッセージから，話題に関する内容を含む重要な文を抽出し，リスト形式で表示する．この抽出方法は，メッセージ中の挨拶文や署名などの不要な文を除き，“>” のマークにより得られる複数メッセージの引用関係を使ってメッセージを連結して，残った文を抽出する．

これらの手法は，予め決めておいた特定のキーワードや記号などを手がかりとして情報を抽出または削除している．そのため，抽出するまたは削除する情報の内容も固定されている．それに対し本手法では，落札者ごとに出品者評価コメントを調べ，各落札者が頻繁に記述する記述を取り出す．そして，落札者ごとに，それらの記述と要約対象の出品者の評価コメントを比較することによって，要約対象の出品者にのみ存在する，または要約対象の出品者にのみ存在しない記述を抽出している．このため，本手法はこれまでの手法と異なり，抽出する情報を出品者と落札者によって適応的に選択していると言える．

表 2 属性と属性値の組の出現頻度
Table 2 The frequency of the pairs of feature and feature value.

属性	属性値	凡例	出現頻度	属性別出現頻度
お礼		～ありがとうございました	156	156
挨拶		またよろしくお願いします	69	69
謝罪		評価が遅れてすみません	12	12
到着	商品	商品受け取りました	34	
	無事	無事に届きました	31	
	本日	本日商品が届きました	28	
	今日	今日受領しました	7	
	すぐ	すぐに届きました	6	
	その他	ちゃんと来ました	13	119
対応	迅速	迅速な対応…	49	
	丁寧	対応も丁寧で…	19	
	親切	親切に対応しださり…	7	
	スムーズ	スムーズな対応で…	3	
	素早い	素早い対応で…	3	
	その他	対応が敏速で…	12	93
取引	スムーズ	スムーズな取引で…	18	
	安心	安心して取引でき…	12	
	気持ちよい	気持ちよく取引でき…	9	
	迅速	迅速に取引してくださり…	5	
	良い	良い取引ができ…	3	
	その他	満足の行く取引ができ…	4	51
発送	迅速	迅速な発送で…	5	
	その他	安く送っていただき…	10	15
連絡	その他	迅速な連絡で…	14	14
梱包	丁寧な	梱包が丁寧で…	8	
	その他	梱包も良好で…	4	12
出品者	信頼	信頼できる出品者です	4	
	その他	信用できる出品者でした	2	6
入金	その他	昨日入金しました	3	3
要望・プラス α	その他	無理な要望を聞いてくれて…	5	5
商品	満足	商品にも満足して…	11	
	良い	商品の状態も良く…	9	
	きれい	とてもきれいで…	5	
	気に入る	とても気に入って…	3	
	大切	大切にします	3	
	その他	貴重なものを…	30	61
その他			19	19
合計			635	635

3. 出品者評価コメントの調査

出品者評価コメントを要約するためには，評価コメントをどのように機械処理するかを決め，効率的に要約するための手がかりを見つける必要がある．そこで，実際の出品者評価コメントに関する調査を行った．3.1 節では，出品者評価コメントの記述内容の調査と文法的な特徴の調査について述べる．記述内容の調査は，出品者評価コメントにはどのような内容が書かれているのかを調べるために行った．また，文法的な特徴の調査は，出品者評価コメントが機械処理できるものなのかどうかを調べるために行った．3.2 節では，出品者評価コメント中に出現するパターンの頻度の調査について述べる．この調査は，自然言語処理において最も一般的な手法である頻度に基づく手法が使えるか否かを検討するために行った．また，3.3 節では，落札者ごとに出品者評価コメントを取り出したときに評価コメント中に出現するパターンの頻度の調査について述べる．そして，3.4 節では，調査結果から得られた結論について述べ，その結論から考えられる要約方法の方針について述べる．

3.1 内容と文法的特徴に関する調査

ランダムに選んだ 18 人の出品者の出品者評価コメントからそれぞれ新しい順に 10 個の評価コメントを取り出した．この合計 180 個の評価コメントの中身が何に関して

記述されているのかを調べた．その結果，出品者の「対応」ややり取りされた「商品」などについて，大きくは13種類のトピックに分類することができた（表2の第一列参照）．そして，それぞれの文には，これらのトピックに対してそれがどのようであったかが記述されている．これらのトピックに対する内容がどういうキーワードで記述されているかを調べてみると，非常に様々なキーワードが存在することが分かった．例えば,「対応」というトピックに関しては,「迅速」や「丁寧」,「素早い」といったキーワードが存在し,「商品」に関しては,「きれい」や「気に入る」といったキーワードが存在する．そこで,「対応」や「商品」などのトピックを属性と定義し,「迅速」や「きれい」などの，属性がどのようであったかを表すキーワードを属性値と定義する．ただし,「ありがとうございました」や「またよろしく」,「すみません」といったお礼や挨拶，謝罪を述べる記述は，上述のような，トピックについてそれがどうであったという記述とならないことがほとんどである．よって，お礼や挨拶，謝罪を述べる記述は例外として扱い，属性を「お礼」や「挨拶」,「謝罪」として属性値はなしとする．このように定義した属性と属性値によって出品者評価コメントの記述内容をおおむね表すことができる．

次に，上述の180個の出品者評価コメントの各文を句で区切ってそれぞれを属性と属性値で表し，属性と属性値の文法的な特徴を調べた．属性の品詞と属性値の品詞の組にどのような組合せがあるかを調べると，24個の組が存在した．この結果を，各組の頻度と共に表3に示す．この表3では，頻度が1である組は「その他」としてまとめて示している．また，属性または属性値に該当する単語が存在しない場合は，属性または属性値をなしとして「ー」で示している．属性が存在しない場合の例は「とてもかわいいです.」や「気に入ってます.」といった文であり，これらは主に商品に関する記述であるが，そうであることを示す属性のキーワードが存在しない．また，属性値が存在しない場合の例は「届きました.」や「到着しました.」といった文であり，これは商品の到着に関する記述であるが属性値が存在しない．表3において,「お礼」,「挨拶」や「謝罪」を除いて1%以上の割合を占める組に注目すると，そのような組は11組存在し，全体の95.4%を占めている．よって，属性の品詞と属性値の品詞の組のほとんどが，この11組のいずれかになると言える．属性として抽出すべき単語を辞書として予め登録しておけば，その単語を探すことで，評価コメント中の属性の位置を特定でき，その属性の品詞と上述の11組の品詞の組から，属性値を抽出することができると思われる．このことから，出品者評価コメントの要約は，上述のように機械処理することで作成可能と思われる．

3.2 内容の頻度に関する調査

上述の180個の出品者評価コメント中の全ての句を，属性と属性値の組によって分類してみると，112種類に分類でき，全部で635個の句が存在した．この分類結果に，属性と属性値の各組に対する凡例と出現頻度を加えて表2に示す．表2を出現頻度に注目して見ると,「お礼」,「挨拶」,「対応-迅速」の組,「到着-商品」の組などにおける出現頻度が高いことが分かる．これらは，出品者評価コメントでは非常に典型的な文章であり，挨拶や儀礼的な一言，決まり文句といった類のものである．そのため，必ずしも出品者や商品の性質を表しているとは限らない．このことから，複数の出品者の比較を目的とする場合，出現頻度の高いこれらの記述から出品者の特徴を得るのは困難と言える．したがって，要約を作成する際にはこれらの出現頻度の高い記述を削除する方が良いと考える．

表3 属性と属性値の品詞の組の出現頻度
Table 3 The frequency of the word classes in the pairs of feature and feature value.

属性の品詞	属性値の品詞	頻度		割合(%)	
お礼	ー	156	606	24.6	95.4
名詞	形容動詞	105		16.5	
動詞	名詞	91		14.3	
挨拶	ー	69		10.9	
名詞	名詞	37		5.8	
動詞	形容動詞	35		5.5	
名詞	形容詞	28		4.4	
名詞	動詞	23		3.6	
ー	動詞	16		2.5	
謝罪	ー	12		1.9	
ー	形容詞	10		1.6	
ー	形容動詞	9		1.4	
ー	名詞	8		1.3	
動詞	副詞	7		1.1	
動詞	形容詞	6	29	0.9	4.6
動詞	ー	3		0.5	
ー	形容詞+名詞	3		0.5	
ー	動詞+名詞	3		0.5	
その他		14		2.2	
合計		635	635	100	100

コメント:今日届きました。無事起動することができました。どうもありがとうございました。

コメント:とても欲しかったのでうれしいです。どうもありがとうございました。

コメント:今日届きました。息子が喜んでいます。どうもありがとうございました。

コメント:今日届きました。新聞なども綺麗で満足しました。どうもありがとうございました。

図1 落札者 A の出品者評価コメントの例
Fig. 1 Examples of feedback comment written by Bidder A.

3.3 落札者ごとに見た内容の頻度に関する調査

10人の落札者が記述した評価コメントのうち，最近のものからそれぞれ10個を取り出した．10人のうち落札者Aの記述した出品者評価コメントの一部を図1に示す．図1の評価コメントを見ると,「どうもありがとうございました」や「今日届きました」などの同じ表現が繰り返されていることが分かる．

3.3節の調査と同様に評価コメントにおける属性と属性値の組の出現頻度を，落札者ごとに求めた．この結果を落札者ごとに分けて表4に示す．ただし，ここでは5人分のみを示す．表4では，どの落札者も特定の属性と属性値の組に関する出現頻度が高くなっている（表4の色の付いたセル）ことが分かる．このことから，落札者は複数の出品者に対して同じような記述をすることが多いと言える．

これらのことから，落札者が出品者についての評価コメントを書く時には，多くの場合いつも決まった内容を記述していることが分かる．このことから，ある特定の

表 4 落札者ごとに調べた属性と属性値の組の出現頻度
Table 4 The frequency of the pairs of feature and feature value in each bidder.

		落札者				
属性	属性値	A	B	C	D	E
お礼		10	2	24	11	10
挨拶		0	10	9	10	4
到着	単に	7	2	2	0	10
	無事に	0	8	7	0	0
取引	スムーズな	0	1	2	0	2
	気持ちよい	0	0	0	0	2
	敏速な	0	0	0	0	1
対応	迅速な	0	0	10	0	1
	素早い	0	0	0	10	2
	丁寧な	0	0	8	0	1
梱包	丁寧な	0	0	1	0	0
	きれい	1	0	0	0	0
商品	気に入る	0	6	0	0	0
	素敵な	0	6	0	0	0
	異常なし	1	0	0	0	0
	良い	0	1	0	0	0
	きれい	1	0	0	0	0
	喜んでいる	3	0	0	0	0
	かわいい	2	0	0	0	0
	安い	0	1	0	0	0
	欲しかった	0	0	0	0	0
出品者	誠実な	0	1	0	0	0
プラス α	おまけ	0	1	0	1	0

落札者が複数の出品者に対して記述した評価コメントのうち，ある出品者にのみ特別に出てくるような記述があったり，逆にある出品者にのみ出てこないような記述があると，落札者はその出品者に対して特別に思い入れを持ったのではないかと推測される．そこで，そのような思い入れを出品者の特徴として抽出するには，出品者ごとに評価コメントを比較するよりも落札者ごとに評価コメントを比較する方が良いと考えられる．

3.4 調査結果から得た結論と要約の方針

3.1 節の調査結果から，出品者評価コメントはそのほとんどが，限られたトピックについてそれがどうであったかということを記述しており，その内容は属性と属性値の組で表せることが分かった．また，その組が単語の品詞情報を使って，取り出せそうなことが分かった．3.2 節の調査結果からは，出現頻度の高い記述は儀礼的な記述であり，要約を作成する際には頻度を用いてこれらの記述を省くべきであるという結論を得た．また，3.3 節の調査からは，出品者評価コメントを落札者ごとに調べる方が儀礼的な記述と有益な記述を判断しやすく，洗練された要約を作成できるという結論を得た．本研究では，これらの結論に基づき，要約を作成する手法である Social Summarization 法（以下，SS 法）を提案する．SS 法では，出品者評価コメントを落札者ごとに調べて，多くの出品者に対して記述されている出現頻度の高い記述を除き，要約対象の出品者にのみ存在する記述を評価コメントの要約とする．また，多くの出品者に記述されているが要約対象の出品者には記述されていないものがあれば，その事実もユーザにとっては有益な情報であると考え，要約対象の出品者にのみ存在しない記述も要約とする．これらの要約を，順に存在要約，非存在要約と呼ぶ．また，SS 法では，記述内容の出現頻度の計算を，属性と属性値に対して行い，精度の高い要約を目指すこととする．

4. Social Summarization 法

SS 法の処理の概要について 4.1 節で述べる．要約対象の出品者とその他の出品者の評価コメントの相違点（要約対象の出品者にのみ存在する記述と要約対象の出品者にのみ存在しない記述）を発見する方法について，4.2 節で述べる．その際に必要な属性と属性値の抽出方法について 4.3 節で述べる．

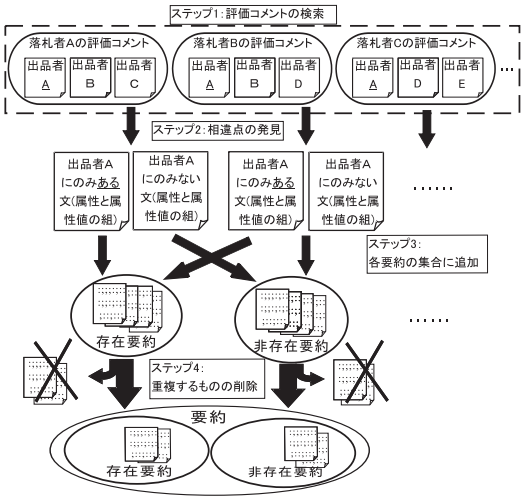


図 2 Social Summarization 法の手順
Fig. 2 Process flow of Social Summarization method.

4.1 SS 法の処理の概要

SS 法によって出品者 A の評価コメントの要約を作成するときの手順を図 2 に示し，以下に SS 法の手順を述べる．

ステップ 1 評価コメントの検索

要約対象の出品者を評価したある落札者が記述した全ての評価コメントを検索する（対象のオークションサイトに関するリンク構造と，ページ構造に関するテンプレートを人手で生成しておき，そのテンプレートに基づき検索する）．

ステップ 2 相違点の発見

要約対象の出品者に対する評価コメントとその他の出品者に対するものを比較して，相違点を見つけ取り出す．この相違点とは要約対象の出品者にのみある記述か，または要約対象の出品者にのみない記述である（これらの記述は実際には，属性と属性値の組で表される）．相違点を見つける方法については，4.2 節で詳しく述べる．

ステップ 3 要約の集合に追加

要約対象の出品者にのみある記述を集めて存在要約と呼ぶ集合に加え，要約対象の出品者にのみない記述を集めて非存在要約と呼ぶ集合に加える．

ステップ 4 重複するものを削除

ステップ 2 とステップ 3 の作業を複数の落札者に対して行い，要約の各集合中に重複して存在するものを削除する．

4.2 相違点の発見方法

ある落札者の評価コメントにおいて，要約対象の出品者とその他の出品者の評価コメントの相違点を見つけるための処理の流れを以下に述べる．

(1) 落札者が記述した評価コメントのうち，要約対象の

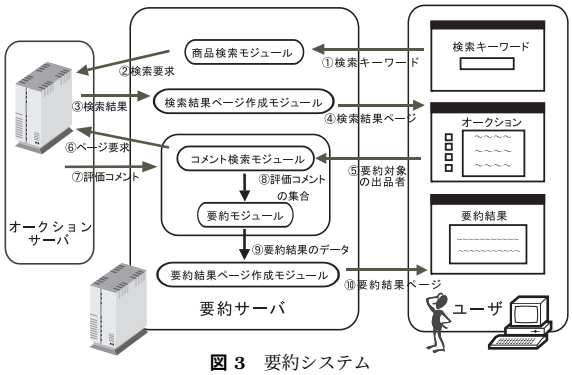


図 3 要約システム
Fig. 3 System structure of our system.

出品者に対する評価コメントを除くその他の評価コメント中に存在するすべての記述（実際には，属性と属性値の組で，その抽出方法は 4.3 節で詳しく述べる）の出現頻度を求める．出現頻度の高い記述（閾値 α 以上）の集合を集合 S とする．

(2) 以下の 2 種類の相違点を取り出す．

- 要約対象の出品者の評価コメント中に存在する記述のうち，集合 S に存在しないものを見つけ取り出す．
- 集合 S の各記述で，要約対象の出品者の評価コメント中には存在しないものを見つけ取り出す．

ここで，前者が要約対象の出品者にのみ存在する記述で，後者が要約対象の出品者にのみ存在しない記述となる．

4.3 属性と属性値の抽出方法

3.1 節の調査に基づいて考案した属性と属性値の抽出方法について述べる．3.1 節の調査では属性の数は限られていることが分かったので，属性を特定するには，予め定義しておいたキーワードを用いることとする．そして，このキーワードと評価コメント中の単語とのキーワードマッチングを行い，マッチしたキーワードを属性と決定する．また，3.1 節の調査では，属性の品詞が決定すれば，おのずと対応する属性値の品詞が決まることが明らかになった．そこで，属性値を特定するには，特定した属性の品詞から対応する属性値の品詞を決定し，その品詞である単語でその属性の位置から最も近い位置にあるものを属性値とすることとする．予め定義しておく属性のキーワードは，3.1 節で得られた 13 種類の属性の類義語とし，類義語辞書で調べたものとした．類義語は各トピックごとに 5 個 ~56 個ある．

予め定義しておく属性値の品詞は，3.1 節で調べた属性と属性値の品詞の組（表 3）の頻度が 1% 以上のものを使う．属性が存在しない組は商品に関する組であることがほとんどであったため，その属性値は商品に関するものと決めている．

5. 要約システム

ユーザが注目している出品者の評価コメントを SS 法を用いて要約し，その要約をユーザに提示する要約システムを実装した（図 3）．この要約システムは，クライアント・サーバ型で実装している（以下，このサーバを要約サーバと呼ぶ）．ユーザに要約を提示するまでの流れを以下に述べる．ユーザは，商品を検索するために検索キーワード入力ページ（図 4）で検索キーワードを入力する．商品検索モジュールがそのキーワードを受け取り，HTTP リクエストをオークションサーバに送る．オークションサーバが検索結果の HTML ソースを検索結果ページ作成モジュールに送る．検索結果ページ作成モジュールがこの HTML ソースから検索結果ページ（図 5）を作成し，ユーザに提示する．ユーザはそのページでオークションのタイトル部分をクリックし，各オークションの詳細情報を閲覧する．そして要約を閲覧したい出品者を選択し，チェックボックスにチェックを入れる．評価コメント検索モジュールは，選択された出品者の評価コメントのページから，要約をするのに必要な評価コメントの検索を行い，検索した全評価コメントを要約モジュールに渡す．その後，要約モジュールが SS 法を用いて評価コメントを要約し，要約結果のデータを要約結果ページ作成モジュールに渡す．要約結果ページ作成モジュールがこのデータから要約結果ページ（図 6）の HTML ソースを作成し，ユーザに提示する．

この要約システムを用いて，Yahoo オークションのある出品者の出品者評価コメントを要約した．要約結果を図 6 に示す．また，要約対象の出品者が記述したオリジナルの評価コメントの一部を図 7（a）に示す．この評価コメントをこのまま要約しようとすると，出品者の特徴を表す有益な記述がどれであるかがわかりにくく，要約することができない．そこで，図 7（a）において○で囲まれた落札者”BBB”に注目する．この落札者が要約対象でない出品者に対して記述した評価コメントの一部を図 7（b）に示す．図 7（a）の”BBB”の記述した評価コメントと図 7（b）の全評価コメントを比較すると，図 7（a）中の”BBB”の記述した評価コメントには「すべて綺麗に再生できました」という記述が存在するが，図 7（b）中には存在しない．よって，この記述は要約対象の出品者にのみ記述された記述であり，図 6 の要約結果に存在要約として現れている．このようにして特別に思い入れを込めて記述された文は，要約として残ることがわかる．

6. 評　　価

本研究における評価の焦点として，まずネットオークションの評価コメントを要約すること自体がユーザが入札する出品者を決定するに当たって有用か否かを検証する．これは，SS 法によって要約した評価コメントを見せる場合と，オリジナルの評価コメントを全て見せる場合で，意思決定にかかった時間を比較することで行う．詳しくは，6.1 節で述べる．次に，SS 法の特徴は，出品者評価コメントを要約するのに，落札者ごとに評価コメントの出現頻度を求める点にある．この特徴の有効性を検証するために，評価コメントを落札者ごと見ずに，出品者ごとに属性と属性値の組の出現頻度を求める手法と比

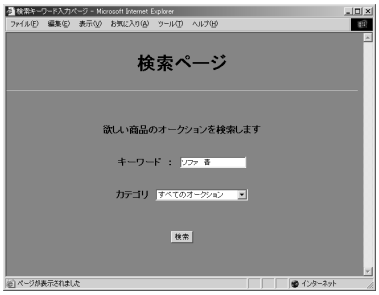


図 4 検索キーワード入力ページ
Fig. 4 A screenshot of the top page.



図 5 検索結果ページ
Fig. 5 A screenshot of the search results page.

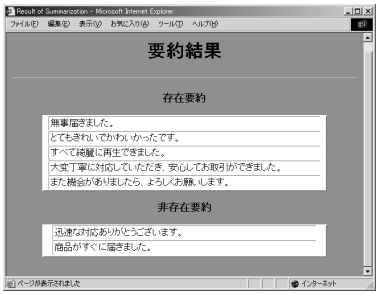
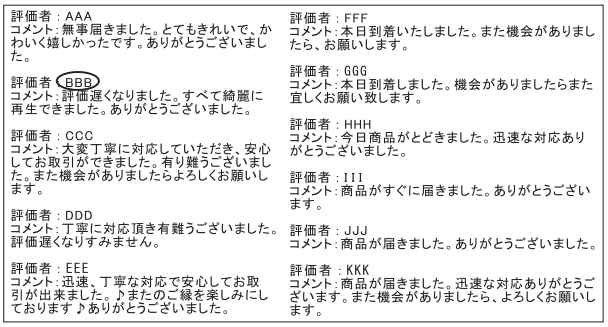
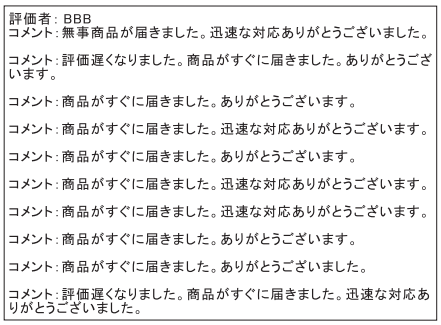


図 6 要約例
Fig. 6 A screenshot of summarization.



（a）要約対象の出品者の評価コメント



（b）他の出品者に対する評価コメント

図 7 オリジナルの評価コメント
Fig. 7 Examples of original feedback comment.

較する．この比較手法自体も過去に提案されたものではないが，その基本的な考え方は，単語に重みを付ける方法である tf・idf に類似し，自然な発想に基づくものと言える．詳しくは，6.2 節，6.3 節で述べる．最後に，6.4 節で評価のまとめを述べる．

6.1 要約することの有用性の検証

ネットオークションの評価コメントを見たことがある学生 4 人を被験者とした．各被験者に，同じカテゴリに出品している 3 人の出品者を比較して，一番良いと思う出品者を一人選択するタスクを行ってもらい，出品者を決定するまでに要した時間を計測する．被験者には，オリジナルの評価コメントを参考にする場合と，SS 法を用いて作成した要約を参考にする場合の 2 種類を順に行ってもらう．ただし，タスクを行うにつれて被験者がタスクに慣れ，所要時間が短くなると考えられるので，被験者を 2 つに分けて実験の順番を変える．また，取引する商品によって評価コメントの書かれ方が異なるため，評価コメントが丁寧に記述される傾向にある「ベビー服」を取引している出品者と，比較的簡素に記述される傾向にある「パソコン」を取引している出品者の両方を対象とする．オリジナルの評価コメントと要約のどちらを参考にする場合も，パソコンを出品している出品者に対してと，ベビー服を出品している出品者に対しての 2 回ずつ，合計 4 回行ってもらう．

ベビー服とパソコンの場合の所要時間の平均値を被験者ごとに求め，表 5 に示す．いずれの被験者も要約を使ったほうが早く出品者を選択できている．被験者数が少ないため，要約を用いる方が良いと結論づけることは現段階ではできないが，被験者を増やしていけば，有意差があるものと思われる．

表 5 所要時間
Table 5 The time required for tasks.

被験者	A	B	C	D	平均
オリジナルの評価コメント	370.0	243.0	213.0	557.5	345.9
SS 法による要約	236.0	133.5	204.5	342.5	229.1

6.2 他手法との比較

6.2.1 節でこの比較手法の処理の流れを説明し，6.2.2 節で，SS 法と比較手法によってできる存在要約の違いについて説明する．

6.2.1 比 較 手 法

SS 法の最大の特徴は，出品者の評価コメントを要約するのに，評価コメントを直接要約するのではなく，その出品者に対してコメントを記述した落札者が他の出品者に対して記述したコメントを利用して要約することにある．これを評価するに当たって，比較手法として，評価コメントを記述した落札者が他の出品者に対して記述したコメントを用いるのではなく，そのコメントが他の出品者にあるか否かを見る方法を採り上げそれと比較する．SS 法との違いは，4.2 節の集合 S を求める際，注目する出品者を除く全ての出品者の評価コメント中に出現する属

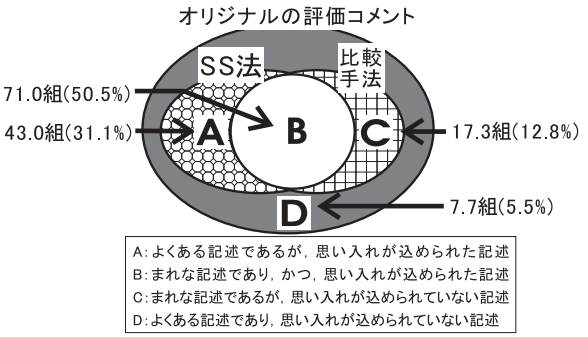


図 8 SS 法と比較手法の関係図
Fig. 8 The difference between the result of Social Summarization method and that of the comparative method.

- とてもかわいいです。
- 早速使わせていただきます。
- 動作確認も問題ありません。
- とても参考になります。
- 大変満足しています。

図 9 比較手法でのみ得られる要約の例
Fig. 9 Examples of the sentence which appears only in the result of Social Summarization method.

性と属性値の出現頻度を求め，その中から出現頻度の高い属性と属性値の組（閾値 α 以上）を取り出す点にある．

6.2.2 存在要約の違い

6 人の出品者を選択して，その 6 人の評価コメントを SS 法，比較手法によって要約する．そして，存在要約中に含まれる属性と属性値の組の異なり数を数えた．また，SS 法と比較手法で共通して存在した属性と属性値の組の数を数えた．6 人の平均値を求め，図 8 に示す．SS 法の存在要約には，ある落札者が他の出品者に対しては書かず，注目する出品者に対してのみ書いた記述が含まれている．つまり，落札者の思い入れが込められた記述が含まれている．また，比較手法の存在要約には，注目する出品者以外の出品者評価コメントには存在しない記述が含まれている．つまり，一般的に見てまれな記述が含まれる．よって，図 8 中の A,B,C,D に含まれる記述は図 8 下部に示す特徴をそれぞれ持っていると言える．

B にはまれであり，かつ思い入れが込められた記述が含まれているので，出品者を選ぶ際に最も注目すべき記述であると言える．また C には，まれであるが，思い入れが込められておらず，ユーザが一見すると重要なコメントとして間違って解釈してしまう非常に問題となる記述が含まれている．例を図 9 に示す．このように一見本心のように見えるが落札者が使い回している記述を，SS 法では除外することができる．一方 A には，よくあるが，思い入れが込められているので参考になると思われる記述が含まれている．A にある記述が出品者選択の判断材料となるか否かが，SS 法の特徴である「コメントを落札者ごとに見る」ことの有効性を検証する鍵となる．

6.3 SS 法でのみ得られる要約の有効性の検証

本節では，SS 法でのみ得られる要約（図 8 の A）が，出品者を比較する際に有効であるか否かを検証する．SS 法でのみ得られる要約をひとつずつ注目して見てみると，それ自体はよく書かれる記述なので，一見すると重要そうに見えず，直接有効性を評価することができない．しかし，そのような記述であっても，その記述に対応する属性と属性値の組の数をカウントすれば（図 2 のステップ 4 で重複を消す前の段階での個数をカウントすれば），どれほどの落札者が本心からそのように思ってその記述を書いたのかが数字として現れると思われる．SS 法による要約において，要約として残すすべての記述の属性と属性値の組の数に対する，注目する記述の属性と属性値の組の割合を求め，さらにオリジナルの評価コメントにおいて，すべての記述の属性と属性値の組の数に対する，注目する記述の属性と属性値 s の組の割合を求め，それらを比較すれば出品者の特徴が得られると思われる．仮説としては，オリジナルの評価コメントには注目する記述であっても，儀礼的に書いたものが多く含まれるため，出品者間のその記述の割合の分散（標準偏差）は小さくなると思われる．一方，SS 法ではそのような儀礼的な記述が除去されるため，本心で書いた記述だけが残り，出品者間のその記述の割合の分散（標準偏差）は大きくなるものと思われる．6.3.1 節で，全体の評価として割合の分散（標準偏差）の平均を用い，オリジナルの評価コメントと SS 法による要約で比較する．6.3.2 節では，割合を表として表示した時に，本当に出品者間で違いが表れるのか否かについて述べる．また，このような違いは，SS 法による要約の割合を見なくてもオリジナルの評価コメントの割合を見るだけで十分にわかる可能性もある．そこで，6.3.3 節では，SS 法による要約とオリジナルの評価コメントで出品者の評価の順番に違いが出てくるか否かを確かめる．

6.3.1 出品者の比較のしやすさ

SS 法でのみ得られる要約中にある属性と属性値の組に注目し，その組がオリジナルの評価コメントに占める割合および要約に占める割合を求めた．そして，組ごとに，割合の標準偏差を，オリジナルの評価コメントの場合と要約の場合でそれぞれ求めた．要約における組の割合の標準偏差が大きい組を取り出し，それらの要約における割合とオリジナルの評価コメントにおける割合を，表 1（a），表 1（b）に分けて示す．オリジナルの評価コメントでの標準偏差と要約での標準偏差を組ごとに比較すると，69 組のうち 67 組で，要約における割合の方がオリジナルの評価コメントにおける割合よりも大きくなった．また，オリジナルの評価コメントと要約で標準偏差の平均を求めたところ，オリジナルの評価コメントでは 0.4 となり，要約では 1.0 となった．このことより仮説が証明された．

6.3.2 出品者の違い

図 10 では，列挙した組の中でも特に注目すべき組の行に薄く影を入れている．図 10 を横に見ると，各組に対し

属性	属性値	割合(%)						標準偏差
		出品者1	出品者2	出品者3	出品者4	出品者5	出品者6	
対応	迅速だ	5.8	6.2	7.2	4	6	8.7	1.4
届く	商品	8.1	11	7.2	7.1	10.6	7.1	1.7
商品	気に入る	0.9	2.3	3.5	0.3	0.7	0	1.2
届く	無事	4.3	4.2	3.5	2.3	1	1.6	1.3
商品	動作	0	0	0	2.3	0	2.2	1.1
挨拶		13.5	12.1	11.8	12.7	13.6	12.1	0.7
届く	本日	3.2	1.4	3.5	0.8	3.7	2.2	1.1
連絡	遅い	1.2	0.8	0.3	0	0	0	0.5
取引	安心	1.2	0.8	2.6	2.8	3	2.8	0.9
商品	確認	0	0	0	1.4	0	1.6	0.7
商品	する	2.9	2	3.5	3.4	2	1.9	0.7
取引	良い	0.3	1.1	0.3	0.8	0	1.6	0.5
届く	すぐだ	0	0.6	0.3	2.3	1	0.6	0.7
対応	丁寧だ	3.5	5.4	2.9	4.5	4	3.7	0.8
謝罪		2	2.3	2.3	1.7	1	1.2	0.5
梱包	丁寧だ	1.4	0.6	0.6	2	1	1.9	0.6
対応	迅速	1.7	1.7	0.9	1.7	1	0.9	0.4
取引	迅速	0.6	0.3	1.2	0.3	1	0.9	0.3

(a) オリジナルの評価コメントにおける割合

属性	属性値	割合(%)						標準偏差
		出品者1	出品者2	出品者3	出品者4	出品者5	出品者6	
対応	迅速だ	11.6	7.1	3.4	3.6	9.3	14.4	4.0
届く	商品	11.6	2	6	4.3	6.2	5.1	2.9
商品	気に入る	0	4	7.7	0.7	1	0	2.8
届く	無事	7	6.1	6.8	3.6	1	0.8	2.6
商品	動作	0	0	0	5	0	5.9	2.6
挨拶		7	2	6.8	5	9.3	2.5	2.6
届く	本日	7	3	3.4	1.4	8.2	3.4	2.4
連絡	遅い	4.7	3	0.9	0	0	0	1.8
取引	安心	1.2	1	2.6	4.3	4.1	5.9	1.8
商品	確認	0	0	0	2.9	0	4.2	1.7
商品	する	2.3	5.1	4.3	5	2.1	1.7	1.4
取引	良い	0	3	0.9	0.7	0	3.4	1.4
届く	すぐだ	0	1	0	2.9	3.1	0	1.3
対応	丁寧だ	3.5	6.1	3.4	6.5	5.2	3.4	1.3
謝罪		2.3	6.1	3.4	2.9	3.1	2.5	1.3
梱包	丁寧だ	2.3	2	1.7	5	2.1	4.2	1.2
対応	迅速	1.2	2	0.9	3.6	1	0	1.1
取引	迅速	1.2	0	3.4	0.7	2.1	0.8	1.1

(b) 要約における割合

図 10 属性値と属性値の組の割合
Fig. 10 The percentage of the pairs of feature and feature value.

て出品者を比較することができる．例えば，1 行目では「対応，迅速だ」の組は，出品者 6 で一番大きな値となっているので，6 人の中で対応が迅速なのは出品者 6 であると推測できる．また，表を縦に見ると，各出品者に特徴的な組が何かを知ることができる．例えば，出品者 6 は「対応が迅速」で「安心して良い取引ができる」と推測でき，出品者 4 は,「商品がすぐ届き」,「対応が迅速かつ丁寧」で,「梱包が丁寧」であると推測できる．このように要約における割合を表にして見ることで，出品者間で違いがあることがわかる．

6.3.3 オリジナルの評価コメントとの違い

図 10 の（a）と（b）を組ごとに比較すると，各組で最も特徴的な出品者がオリジナルの評価コメントと要約で異なる組がいくつか見られる．ほとんどの組では同じであるが,「対応，迅速だ」の組に注目して見ると，オリジナルの評価コメントでは出品者 3 が 7.2 で 2 番目に特徴的な出品者であるが，要約では出品者 3 は 3.4 で 6 番目に特徴的な出品者となっている．ユーザがオリジナルの評価コメントで「対応，迅速だ」の組の値を見ると,「出品者 3 の対応は迅速である」と解釈するが，要約で見ると「出品者 3 の対応はそこまで迅速でない」と解釈する．このようにオリジナルの評価コメントでは儀礼的な記述がノイズとなり，出品者の評価の順番が変わってくることがわかる．

6.4 評価のまとめ

6.1 節の被験者実験から，要約すること自体がユーザが入札する出品者を決定するに当たって有用であることがわかった．6.2 節での他手法との比較からは，SS 法と他手法には要約の内容に差があることがわかった．SS 法は，比較手法で得られるまれな記述のうち，本心から述べられていない記述を削除でき，よくある記述でも本心から述べられているものを残すことがわかった．6.3 節の検証からは，SS 法でのみ得られる要約も出品者の特徴を表しており，入札する出品者を決定する際に参考にすることができるとわかった．以上の結果より，今後は，SS 法と他手法で残る要約の違いを考慮した要約の提示方法を考える必要がある．

7. まとめと今後の課題

本研究では，出品者評価コメントを要約する際に評価コメントをどのように機械処理するかを決めるために，また効率的に要約するための手がかりを見つけるために，出品者評価コメントを調査した．その調査から，出現頻度の高い記述を省くべきであり，その際には，出品者評価コメントを落札者ごとに調べる方が良いという結論を得た．この結論に基づいて，Social Summarization 法（以下，SS 法）という要約の手法を提案した．SS 法は，出品者評価コメントを落札者ごとに調べて，要約対象の出品者にのみ存在する記述，または要約対象の出品者にのみ存在しない記述を抽出する手法である．この SS 法を用いて評価コメントを要約し，その要約をユーザに提示する要約システムを実装した．そして，入札する出品者を決定する際に要約すること自体が有効であるか否か，SS 法の要約が有効であるか否か，SS 法でのみ得られる存在要約が有用であるか否かを検証した．課題として，SS 法の要約をどのように提示するかということが明らかになった．今後は，この課題に取り組むべく，ユーザインタフェースの改良を行っていきたい．

参 考 文 献

1) 乃村 能成，花田 泰紀，牛島 和夫：MHC-Message Harmonized Calendaring System の設計と実装，情報処理学会論文誌，Vol.42，No.10，pp.2518-2525 (2001).
2) 佐藤 円，佐藤 理史，篠田 陽一：電子ニュースのダイジェスト自動生成，情報処理学会論文誌，Vol.36，No.10，pp.2371-2379 (1995).
3) 佐藤 理史，佐藤 円：ネットニュースグループ fj.wanted のダイジェスト自動生成，自然言語処理，Vol.3，No.2，pp.19-32 (1996).
4) 梅木 秀雄：コミュニケーションに埋もれた知識を活用するコミュニティウェア，情報処理学会誌，Vol.43，No.10，pp.1085-1092 (2002).
5) 妙木 裕，松本 裕治，長尾 眞：汎用日本語辞書および形態素解析システム，情報処理学会第 42 回全国大会予稿集 (1991).